Realistic Estimation of Model Parameters

Pavel Ettler
COMPUREG Plzeň, s.r.o., 306 34, Plzeň, Czech Republic

Keywords: Parameter Estimation, Process Model, Probabilistic Distribution, Adaptive Control.

Abstract: Most often, the normal distribution $\mathcal{N}$ plays the key role in the process modelling and parameter estimation. The paper deals with *realistic* estimation of model parameters which takes into account limitations on parameters which arise in industrial applications of the model-based adaptive control. Here the limitation of a normally distributed random variable is being modelled by specific distribution – the probabilistic mixture $\mathcal{D}$. It is shown that relationship between distributions $\mathcal{N}$ and $\mathcal{D}$ coincides with properties of the generalized normal distribution $\mathcal{G}$ and that relations between their first and second statistical moments can be adequately approximated by $\mathcal{G}$'s cumulative distribution function and probability density function, respectively. The derived method is then applied to estimation of bounded parameters. In combination with the idea of parallel identification of the *full* and *reduced* models of the process, a working algorithm is derived. Performance of the algorithm is illustrated by examples on both simulated and real data.

1 INTRODUCTION

The notion of adaptive control first appeared already in the fifties of the last century – see, e.g. (Åström and Kumar, 2014) for further references. Its boom started in the seventies with the expectation of broad utilization of the adaptive control not only in the industrial practice. Presently, this type of control has found its place in real applications but in the extent far from earlier assumption. Among reasons of such progress and today's situation there is one that originates in the mismatch between the process model and the real system or more likely, in disrespecting such discrepancy which always exists in practice. It is closely connected with the fact that measured data which are available for the model and estimation of its parameters are always burdened with uncertainty.

Searching for a model which preferably approximates behaviour of the observed or even controlled system in all situations and handling of every conceivable exceptions turned out to be so demanding and time-consuming that other types of control, PID control in particular, dominate the scene regarding number of applications. In effect, the model-based adaptive control is being employed just in cases where increased demands on its implementation pay off and when the search for a model based on the imperfect data leads to an acceptable result.

It proves that the model-based control is facing the problem which is inherent also for other theoretical approaches to control: if we are "inside" the elaborated theory everything fits together and all tasks seem solvable. The problem occurs on the boundary of the given theory and real world – the reality does not respect prerequisites of the approach, the model is just approximation of the real system, data are corrupted by noise or burdened by uncertainty in general.

The aim of the paper is to contribute usefully to solution of the problem in the borderland between theory and reality.

1.1 Bounded Parameter Estimation

Bounded estimation issues are anything but new and motivation for the solution exists in many application fields which corresponds to the variedness of journals and proceedings in which at least partial solutions are being published. Nevertheless, it has been observed (Murakami and Seborg, 2000; Kopylev, 2012) that very few thorough monographs exist in this respect (Van Eeden, 2006).

1.2 Existing Variety of Solutions

As the result, state of affairs is rather disorganized but even so, the solutions can be divided by the type of limitation being applied (with citation examples in brackets):

- Limitation on the estimation error (Milanese et al., 1996),
- Limitation on the system noise (Norton, 1987),
- Confidence intervals (Mandelkern, 2002),

Ettler, P.
Realistic Estimation of Model Parameters.
DOI: 10.5220/0006395705270534
In *Proceedings of the 14th International Conference on Informatics in Control, Automation and Robotics (ICINCO 2017) - Volume 1*, pages 527-534
ISBN: 978-989-758-263-9

- Statistically bounded (*soft-constrained*) estimation (Benavoli et al., 2006),
- Sharp restrictions on the parameter estimates (Kárný, 1982; Benavoli et al., 2006; Ettler and Kárný, 2010).

The presented solution belongs to the latter type.

2 PROCESS MODEL AND ITS PARAMETERS

Consider a process model for which $X \doteq [y(t), z(t)]$ stand for vector of observations, where the output y is in discrete time instants $t = 1, 2, \ldots$ related to measured data

$$\begin{aligned} z(t) &= [y(t-1),\ \ldots,\ y(t-m_y), \\ &= \quad u_1(t),\ \ldots,\ u_1(t-m_u), \\ &= \quad v(t),\ \ldots,\ v(t-m_v), 1] \end{aligned} \tag{1}$$

Vector z is composed by m_y samples of past outputs y, $m_u + 1$ samples of controllable inputs u and $m_v + 1$ samples of measurable disturbances v. System output y is here considered as scalar for the sake of simplicity, while inputs and disturbances can be multidimensional. Number one as the last vector element enables to consider an absolute term alias offset. The considered stochastic relation is parameterized by unknown parameters Θ with finite dimension and can be described by the probability density function (pdf) $f(y|z,\Theta)$.

Note: *Most of the discussed variables are considered to be time-variant. Nevertheless, the time index is sometimes omitted in the following text for the sake of simplicity*.

2.1 Linear Gaussian Model

In practice, the linear regression model is mostly considered and the uncertainty is approximated by the normal (Gaussian) probability distribution $\mathcal{N}$ defined by two parameters

$$f(y|z,\Theta) = \mathcal{N}(\hat{y}, \sigma_y^2)\ , \tag{2}$$

where $\hat{y}(t) = \theta'(t)z(t)$ and

$$\Theta = [\theta, \sigma_y^2] \tag{3}$$

are unknown parameters of the system,
while $\theta = [\vartheta_1,\ \vartheta_2,\ \ldots,\ \vartheta_m]'$, $m = m_y + (m_u + 1) + (m_v + 1) + 1$ and σ_y^2 stands for the variance of the system noise. Equivalently, the model can be expressed as

$$y(t) = \theta'(t)z(t) + e(t) \tag{4}$$

with the system noise $e(t) \sim \mathcal{N}(0, \sigma_y^2)$.

2.2 Estimator of Unknown Parameters

Author's regular choice regarding parameter estimation is Bayesian probabilistic approach. Here unknown parameters can be considered to be random variables described by the pdf $f(\Theta|y,z)$. It can be derived (Peterka, 1981; Kárný et al., 2005)) that for the model (4) and fixed parameters the pdf is fully specified by the positive definite information matrix V which – after introduction of some type of forgetting (Kulhavý and Zarrop, 1993) to allow tracking of varying parameters – can be recursively updated as

$$V(t) = \varphi V(t-1) + [y(t), z(t)]'[y(t), z(t)], \tag{5}$$

where $\varphi \in \langle 0, 1\rangle$ is forgetting factor. Partitioning of V

$$V = \begin{bmatrix} V_y & V'_{zy} \\ V_{zy} & V_z \end{bmatrix} \tag{6}$$

enables to express the parameter estimates

$$\hat{\theta} = V_z^{-1} V_{zy}\ . \tag{7}$$

Variances of system noise and parameter estimates, respectively, can be estimated as

$$\hat{\sigma}_y^2 = \frac{V_y - V'_{yz} V_z^{-1} V_{zy}}{\kappa}\ , \tag{8}$$

$$\hat{\sigma}_\vartheta^2 = \hat{\sigma}_y^2 \,\mathrm{diag}(\mathrm{V}_z^{-1})\ , \tag{9}$$

where

$$\kappa(t) = \varphi\kappa(t-1) + 1\ . \tag{10}$$

In real-time applications it is appropriate to work with V in its factorized form $V^{-1} = LDL'$ (Peterka, 1981) for the sake of numerical stability.

3 RESPECTING LIMITATIONS OF REAL-LIFE QUANTITIES

Even if unlimited quantities exist in the real world, our observations of them are always bounded. Findings of consideration about an observed general bounded variable will be used for the parameter estimates in the following sections.

3.1 Probabilistic Formulation of Limitation

Consider a real-world variable ξ and its observation x in discrete time instants $t = 0, 1, \ldots$. The observation is bounded by $\underline{x} = x_{\min}$ a $\overline{x} = x_{\max}$ which are given e.g. by a sensor range. Then it holds

$$x(t) = \begin{cases} \underline{x} & \text{for } \xi(t) \le \underline{x} \\ \xi(t) & \text{for } \xi(t) \in (\underline{x}, \overline{x}) \\ \overline{x} & \text{for } \xi(t) \ge \overline{x}\ . \end{cases} \tag{11}$$

Let ξ be considered a normally distributed random variable $\xi \sim \mathcal{N}(\mu_\xi, \sigma_\xi^2)$. Owing to the limitation (11), distribution of x is not Gaussian but can be expressed by a probabilistic mixture $\mathcal{D}$ the pdf of which is composed by a central part of normal pdf and two Dirac functions δ at boundary points of the interval $\langle \underline{x}, \overline{x} \rangle$ (Benavoli et al., 2006).

Then, the pdf of $\mathcal{D}$ is represented by

$$\begin{aligned} f_{\mathcal{D}}(x \,|\, \mu_x, \sigma_x^2, \underline{x}, \overline{x}) \;&=\; A\,\delta(|\xi - \underline{x}|) + \\ &\quad f_{\mathcal{N}}(\xi \in \langle \underline{x}, \overline{x} \rangle \,|\, \mu_\xi, \sigma_\xi^2) + \\ &\quad B\,\delta(|\xi - \overline{x}|)\;, \end{aligned} \tag{12}$$

where

$$A \;=\; F_{\mathcal{N}}(\underline{x} \,|\, \mu_\xi, \sigma_\xi^2) \tag{13}$$

$$B \;=\; 1 - F_{\mathcal{N}}(\overline{x} \,|\, \mu_\xi, \sigma_\xi^2)\;, \tag{14}$$

where $F_{\mathcal{N}}$ stands for cumulative distribution function (cdf) of normal distribution and cdf of $\mathcal{D}$ reads

$$F_{\mathcal{D}}(x \,|\, \mu_x, \sigma_x^2, \underline{x}, \overline{x}) = \begin{cases} 0 & \text{for } \xi < \underline{x} \\ F_{\mathcal{N}}(\xi \,|\, \mu_\xi, \sigma_\xi^2) & \text{for } \xi \in \langle \underline{x}, \overline{x} \rangle \\ 1 & \text{for } \xi > \overline{x}\,. \end{cases} \tag{15}$$

Example in Fig. 1 shows pdf and cdf of $\mathcal{D}$ with parameters $\mu_x = 0.3$, $\sigma_x^2 = 0.7$, $\underline{x} = -1$ a $\overline{x} = 1$.

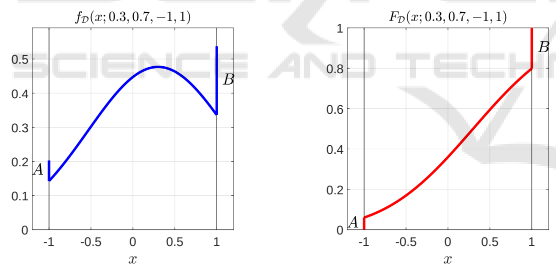

Figure 1: Pdf and cdf of the probabilistic mixture $\mathcal{D}$.

3.2 Distribution $\mathcal{D}$ and Parameter Estimation

Obviously, properties of the normal distribution $\mathcal{N}$ which make it exceptionally suitable for estimation algorithm development cannot be expected in the heterogenous distribution $\mathcal{D}$.

Exact solution can be found only in the case of one bounded parameter (Benavoli et al., 2006). General case leads to utilization of iterative numerical methods which are not very suitable for real-time applications. It is necessary to look for another solution.

3.3 Relations between Statistical Moments of Distributions $\mathcal{N}$ and $\mathcal{D}$

As a consequence of (11), mean values μ_x, μ_ξ and variances σ_x^2, σ_ξ^2 will differ while it holds

$$\mu_x \;\in\; \langle \underline{x}, \overline{x} \rangle, \left\{ \begin{array}{ll} \mu_x > \mu_\xi & \text{for } \mu_\xi < M \\ \mu_x = \mu_\xi & \text{for } \mu_\xi = M \\ \mu_x < \mu_\xi & \text{for } \mu_\xi > M \end{array} \right\} \tag{16}$$

$$M \;=\; (\underline{x} + \overline{x})/2 \tag{17}$$

$$\sigma_x^2 \;\in\; \langle 0, \sigma_\xi^2 \rangle\,. \tag{18}$$

In the case of success to evaluate relation

$$\{\mu_\xi, \sigma_\xi^2, \underline{x}, \overline{x}\} \rightarrow \{\mu_x^\star, \sigma_x^{2\star}\}\;, \tag{19}$$

while respecting conditions (17) and (18), it would be possible to apply the rule for calculation of bounded parameter estimates Θ.

3.4 Employment of the Generalized Normal Distribution $\mathcal{G}$

Extensive experiments led to the result that relations between variances $\sigma_{\mathcal{N}}^2$ and $\sigma_{\mathcal{D}}^2$ of distributions $\mathcal{N}$ and $\mathcal{D}$ can be successfully approximated by the pdf of the generalized normal distribution $\mathcal{G}$ and relations between means $\mu_{\mathcal{N}}$ and $\mu_{\mathcal{D}}$ can be approximated by cdf of the same distribution (see Appendix for properties of distribution $\mathcal{G}$).

Simple rules were found for computation of $\mathcal{G}$'s parameters $\mu_{\mathcal{G}}$, α and β:

$\mu_{\mathcal{G}}$: For the allowable range $x \in \langle \overline{x}, \underline{x} \rangle$, it is natural to place the mean $\mu_{\mathcal{G}}$ in the middle M (17) of the interval.

α: It is possible to define width of the distribution $\mathcal{G}$ as a distance between inflection points of its pdf while these points coincide with limits of the interval $\langle \overline{x}, \underline{x} \rangle$. Then, parameter α can be expressed as $\alpha = (\overline{x} - \underline{x})/2$.

β: Experiments have shown that the shape of $f_{\mathcal{G}}$ remains the same for constant ratio of the interval span (i.e. of parameter α) to the standard deviation σ_ξ. Therefore the shape parameter β can be defined as $\beta = \alpha/\sigma_\xi$.

To sum up, following relations hold for $\underline{x} = 0$, $\overline{x} = 1$ and $\sigma_\xi = 1$:

$$\alpha = \frac{\overline{x} - \underline{x}}{2} \tag{20a}$$

$$\beta = \frac{\alpha}{\sigma_\xi} \tag{20b}$$

$$\mu_{\mathcal{G}} = \underline{x} + \alpha \tag{20c}$$

Modification of equations (20) for general case is described below.

3.5 Limiting Functions ℓ_μ, ℓ_{σ^2}

As was already mentioned, group of equations (20) allows to construct $f_\mathcal{G}$ and $F_\mathcal{G}$ which approximate relationship between unbounded and bounded moments for the special case ($\underline{x} = 0, \overline{x} = 1, \sigma_\xi = 1$).

Let call the limiting functions for both moments ℓ_μ, ℓ_{σ^2}. Following restrictive conditions enabled to find formulation of the functions for general values of boundaries $\underline{x}$, $\overline{x}$ and variance σ_ξ^2:

$$\min_{\mu_\xi \in R} \ell_\mu(\mu_\xi | \sigma_\xi^2, \underline{x}, \overline{x}) = \underline{x} \tag{21a}$$

$$\max_{\mu_\xi \in R} \ell_\mu(\mu_\xi | \sigma_\xi^2, \underline{x}, \overline{x}) = \overline{x} \tag{21b}$$

$$\min_{\mu_\xi \in R} \ell_{\sigma^2}(\mu_\xi | \sigma_\xi^2, \underline{x}, \overline{x}) = 0 \tag{21c}$$

$$\max_{\mu_\xi \in R} \ell_{\sigma^2}(\mu_\xi | \sigma_\xi^2, \underline{x}, \overline{x}) = \sigma_\xi \tag{21d}$$

While $f_\mathcal{G}$ is a non-negative symmetrical function, its maximum lies in the middle $\mu_\mathcal{G}$ (20c) of the bounding interval; therefore it holds for extreme values of $f_\mathcal{G}$

$$f_{\mathcal{G},\min} = 0 \tag{22a}$$

$$f_{\mathcal{G},\max} = f_\mathcal{G}(\mu_\mathcal{G} | \mu_\mathcal{G}, \alpha, \beta) \,. \tag{22b}$$

As $F_\mathcal{G}$ represents cdf it holds

$$F_{\mathcal{G},\min} = F_\mathcal{G}(\xi_{-\infty} | \mu_\mathcal{G}, \alpha, \beta) = 0 \tag{23a}$$

$$F_{\mathcal{G},\max} = F_\mathcal{G}(\xi_{\infty} | \mu_\mathcal{G}, \alpha, \beta) = 1 \tag{23b}$$

The first conversion coefficient can be specified from (21d) and (22b)

$$K_{\sigma^2} = \frac{\sigma_\xi^2}{f_{\mathcal{G},\max}} \tag{24}$$

and the second one follows from (21a), (21b) and (23)

$$K_\mu = \overline{x} - \underline{x} \,. \tag{25}$$

Now the sought limiting functions can be defined as

$$\ell_\mu(\mu_\xi | \sigma_\xi^2, \underline{x}, \overline{x}) = \underline{x} + K_\mu F_\mathcal{G}(\mu_\xi | \mu_\mathcal{G}, \alpha, \beta) \tag{26a}$$

$$\ell_{\sigma^2}(\mu_\xi | \sigma_\xi^2, \underline{x}, \overline{x}) = K_{\sigma^2} f_\mathcal{G}(\mu_\xi | \mu_\mathcal{G}, \alpha, \beta) \,. \tag{26b}$$

The use of the limiting functions and quality of approximation are illustrated in Fig. 2.

4 BOUNDED PARAMETER ESTIMATION

Results obtained for general random variable x can now be engaged for bounding of estimated parameters. It can be formally described as

$$\hat{\theta}^\star = \ell_\mu(\hat{\theta}, \hat{\sigma}_\theta^2, \underline{\theta}, \overline{\theta}) \tag{27a}$$

$$\hat{\sigma}_\theta^{2\star} = \ell_{\sigma^2}(\hat{\theta}, \hat{\sigma}_\theta^2, \underline{\theta}, \overline{\theta}) \,, \tag{27b}$$

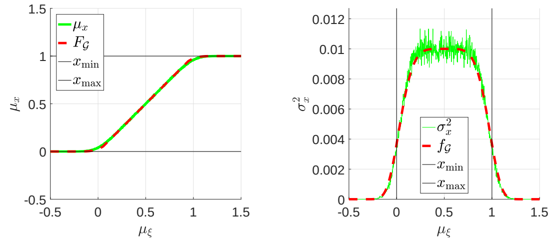

Figure 2: Limitations of the mean and variance of bounded random variable x and their approximation by converted cdf and converted pdf of distribution $\mathcal{G}$. The conversions were realized by the limiting functions ℓ_μ, ℓ_{σ^2}.

where symbol * denotes bounded values of original estimates $\hat{\theta}$, $\hat{\sigma}_\theta^2$ and $\underline{\theta}$, $\overline{\theta}$ are lower and upper parameter boundaries, respectively.

4.1 Special Case: Single Parameter

Consider a single parameter model

$$y(t) = \vartheta u(t) + e(t) \,. \tag{28}$$

In this simple case, information matrix V can be partitioned into scalars according to (6)

$$V = \begin{bmatrix} V_y & V_{zy}' \\ V_{zy} & V_z \end{bmatrix} = \begin{bmatrix} v_y & v_{zy} \\ v_{zy} & v_z \end{bmatrix} \,. \tag{29}$$

Given $\underline{\vartheta}$, $\overline{\vartheta}$ and limiting functions (27) for evaluation of the limited estimate $\hat{\vartheta}^\star$, a modified matrix is sought as

$$V^\star = \begin{bmatrix} v_y^\star & v_{zy}^\star \\ v_{zy}^\star & v_z^\star \end{bmatrix} \,, \tag{30}$$

which corresponds to bounded values $\hat{\vartheta}^\star$, $\hat{\sigma}_\vartheta^{2\star}$. Original estimate can be described regarding to (7) as

$$\hat{\vartheta} = \frac{v_{zy}}{v_z} \,. \tag{31}$$

Using (8) the variance of output reads

$$\hat{\sigma}_y^2 = \frac{v_y - v_{zy}^2 / v_z}{\kappa(t)} \,. \tag{32}$$

Variance of the parameter estimate (9) is

$$\hat{\sigma}_\vartheta^2 = \frac{\hat{\sigma}_y^2}{v_z} = \frac{v_y / v_z - v_{zy}^2 / v_z^2}{\kappa(t)} \,. \tag{33}$$

The estimate of the bounded parameter is then given by

$$\hat{\vartheta}^\star = \frac{v_{zy}^\star}{v_z^\star} = \ell_\mu(\hat{\vartheta}, \hat{\sigma}_\vartheta^2, \underline{\vartheta}, \overline{\vartheta}) \,. \tag{34}$$

As the variance of measured output does not depend on the estimated parameter it can be assumed that

$$v_y^\star = v_y \tag{35}$$

and therefore

$$\hat{\sigma}_{\vartheta}^{2\star} = \frac{v_y/v_z^\star - v_{zy}^{\star 2}/v_z^{\star 2}}{\kappa(t)} \,. \tag{36}$$

It holds (34)

$$\hat{v}_{zy}^\star = \hat{\vartheta}^\star \, \hat{v}_z^\star \,, \tag{37}$$

which using (36) results in

$$\hat{\sigma}_{\vartheta}^{2\star} = \frac{v_y/v_z^\star - \hat{\vartheta}^{\star 2}}{\kappa(t)} = \ell_{\sigma^2}(\hat{\vartheta}, \hat{\sigma}_{\vartheta}^2, \underline{\vartheta}, \overline{\vartheta}) \tag{38}$$

$$\hat{v}_z^\star = \frac{v_y}{\kappa \hat{\sigma}_{\vartheta}^{2\star} + \hat{\vartheta}^{\star 2}} \,. \tag{39}$$

After introducing function ℓ_V which modifies matrix V while using relations (35), (37) and (39), the modified matrix can be written as

$$V^\star = \begin{bmatrix} v_y & v_{zy}^\star \\ v_{zy}^\star & v_z^\star \end{bmatrix} = \ell_V(V, \hat{\vartheta}^\star, \hat{\sigma}_{\vartheta}^{2\star}) \,. \tag{40}$$

Resulting algorithm consists of two parts:

- Initialization

$$V^\star(0) = kI \tag{41a}$$
$$\kappa(0) = 1 \,, \tag{41b}$$

where I is an identity matrix and $k > 0$ is an initialization constant.

- One step of recursion

$$V(t) = \varphi V^\star(t-1) + [y(t), z(t)]' \, [y(t), z(t)] \tag{42a}$$
$$\kappa(t) = \varphi \kappa(t-1) + 1 \tag{42b}$$
$$\hat{\vartheta}(t) = \frac{v_{zy}(t)}{v_z(t)} \tag{42c}$$
$$\hat{\sigma}_y^2(t) = \frac{v_y(t) - v_{zy}^2(t)/v_z(t)}{\kappa(t)} \tag{42d}$$
$$\hat{\vartheta}^\star(t) = \ell_\mu(\hat{\vartheta}(t), \hat{\sigma}_{\vartheta}^2(t), \underline{\vartheta}, \overline{\vartheta}) \tag{42e}$$
$$\hat{\sigma}_{\vartheta}^{2\star}(t) = \ell_{\sigma^2}(\hat{\vartheta}(t), \hat{\sigma}_{\vartheta}^2(t), \underline{\vartheta}, \overline{\vartheta}) \tag{42f}$$
$$V^\star(t) = \ell_V(V(t), \hat{\vartheta}^\star(t), \hat{\sigma}_{\vartheta}^{2\star}(t)) \tag{42g}$$

4.1.1 Illustrative Example

Consider a model

$$y(t) = \vartheta u(t-1) + e_y(t) \tag{43}$$

with one parameter

$$\vartheta = 0.15 \,, \tag{44}$$

where boundaries, noise and number of steps are

$$\begin{array}{rcl} \underline{\vartheta} & = & 0.0 \qquad \overline{\vartheta} \; = \; 0.25 \\ e_y(t) & \sim & \mathcal{N}(0,1) \\ n & = & 1000 \end{array} \tag{45}$$

The control signal was generated according to equation

$$u(t) = \sin(t/20) \,. \tag{46}$$

Behaviour of input $u(t)$ and output $y(t)$ are shown in Fig. 3. Due to intentionally big variance of the system noise, the harmonic component can be hardly recognized in the output signal.

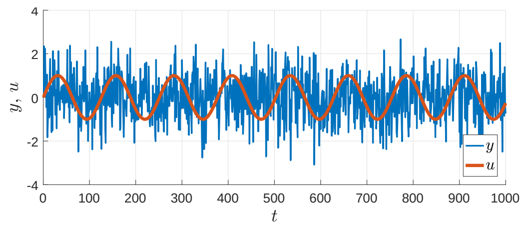

Figure 3: Special case: single parameter model – behaviour of input and output.

The unbounded parameter was estimated by the algorithm described in section 2.2. The bounded estimate resulted from algorithm (42). Behaviour of both unbounded (blue) and bounded (green) parameter estimates is depicted in Fig. 4. Value of the true parameter is represented by the yellow line and the black lines show the given boundaries.

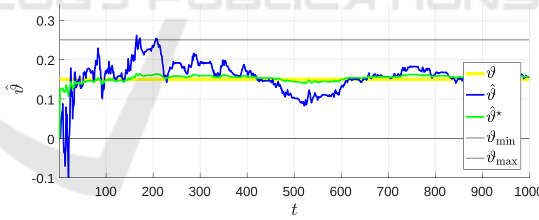

Figure 4: Special case: single parameter model – behaviour of unbounded (blue) and bounded (green) parameter estimates.

4.2 General Case

Exact solution for general case cannot be found because of the absence of a definite rule how to modify an off-diagonal element in a general position within the information matrix. An alternative solution had to be found.

Let the parameters of the general model (4) be divided into two parts

$$\theta = [\theta_a, \theta_b] \,, \tag{47}$$

where θ_a is a m_a-elements parameter vector for which limitation should be applied and θ_b is a m_b-elements vector of parameters without boundaries. It must hold

$$m = m_a + m_b,\ m_a < m,\ m_b \geq 1\ . \qquad (48)$$

Process model which uses all the defined parameters can be called the *full model*

$$y = \theta_a' z_a + \theta_b' z_b + e\ , \qquad (49)$$

where z_a, z_b are parts of the data vector z (2) which correspond to the parts of parameter vector (47).

Unbounded estimation can be based on the algorithm from 2.2 while the vector Θ (3) was enlarged by estimates of parameter variances σ_θ^2

$$\Theta = [\theta, \sigma_\theta^2, \sigma_y^2]\ . \qquad (50)$$

Note: Calculation of parameter variances according to (9) *is influenced by behaviour of the auxiliary variable* κ *defined by* (10)*. It might be better to calculate the variances explicitly in a moving window the length of which enables to tune sensitivity of the algorithm to variance changes.*

4.2.1 Basic Algorithm of Bounded Estimation

The basic algorithm can be divided into following parts:

- Initialization of V and κ similarly to (41)
- One step of the recursion
  - Estimation $\hat{\Theta} = [\hat{\theta}, \hat{\sigma}_\theta^2, \hat{\sigma}_y^2]$ of the full model.
  - Application of the limiting function (27a) on subset of the estimates $\hat{\theta}_a$

    $$\hat{\theta}_a^\star = \ell_\mu(\hat{\theta}_a, \hat{\sigma}_{\theta,a}^2, \underline{\theta}_a, \overline{\theta}_a) \qquad (51)$$

    Vector of estimated parameters $\hat{\theta}^\circ$ is now composed by the bounded and unbounded estimates

    $$\hat{\theta}^\circ = [\hat{\theta}_a^\star, \hat{\theta}_b]\ . \qquad (52)$$

4.2.2 Enlarged Algorithm

Bounded parameter estimates $\hat{\theta}_a^\star$ from the basic algorithm can be, for a particular recursion step, temporarily considered known constants. Then, they can be – together with corresponding data z_a – moved to the left side of model equation. It results in *reduced model*

$$y - \hat{\theta}_a^{\star\prime} z_a = \theta_b' z_b + e\ , \qquad (53)$$

the parameters θ_b' of which can be newly estimated by the common algorithm from 2.2.

Thus, the basic algorithm is enlarged to consist of parts

- Initialization of V and κ similarly to (41)
- One step of the recursion
  - Estimation $\hat{\Theta} = [\hat{\theta}, \hat{\sigma}_\theta^2, \hat{\sigma}_y^2]$ of the full model.
  - Employment of the limiting function (27a) on subset of the estimates $\hat{\theta}_a$

    $$\hat{\theta}_a^\star = \ell_\mu(\hat{\theta}_a, \hat{\sigma}_{\theta,a}^2, \underline{\theta}_a, \overline{\theta}_a) \qquad (54)$$

  - Moving $\hat{\theta}_a^\star$ together with data z_a to the left side to create the reduced model (53)
  - Estimation of parameters $\hat{\theta}_b^\star$ of the reduced model. Vector of estimated parameters $\hat{\theta}^\star$ is now composed by the bounded estimated parameters of the full model and by the modified estimate of unbounded parameters coming from the reduced model

    $$\hat{\theta}^\star = [\hat{\theta}_a^\star, \hat{\theta}_b^\star]\ . \qquad (55)$$

The enlarged algorithm ensures lesser prediction error than the basic algorithm. In many real cases is the prediction error even comparable to the one of the entirely unbounded estimate. Particular results can be influenced by the choice of forgetting factors for estimation of the full and reduced models.

4.2.3 Simulated Example

Consider a process model

$$y(t) = \vartheta_1 y(t-1) + \vartheta_2 u(t-1) + \vartheta_3 + e_y(t) \qquad (56)$$

with parameters

$$\vartheta_1 = 0.8 \qquad \vartheta_2 = 0.3 \qquad \vartheta_3 = 4.0 \qquad (57)$$

and given boundaries, noise and number of samples

$$\begin{array}{rclrcl} \underline{\vartheta}_1 & = & 0.6 & \overline{\vartheta}_1 & = & 0.9 \\ \underline{\vartheta}_2 & = & 0.0 & \overline{\vartheta}_2 & = & 0.35 \\ e_y(t) & \sim & \mathcal{N}(0,1) & & & \\ n & = & 5000 & & & \end{array} \qquad (58)$$

The control signal was generated from

$$u(t) = u(t-1) + e_u(t)\ , \qquad (59)$$

where $e_u(t)$ is a random variable with the uniform distribution $e_u(t) \sim \mathcal{U}(-0.5, 0.5)$.

Behaviour of input $u(t)$ and output $y(t)$ is shown in Fig. 5.

Behaviour of parameter estimates is depicted in Fig. 6. True parameters are represented by yellow lines while black lines in the first two graphs represent their given boundaries. Unbounded estimates are depicted in blue while the bounded ones are green. The estimate of the single-parameter reduced model is plotted in red in the lowermost graph.

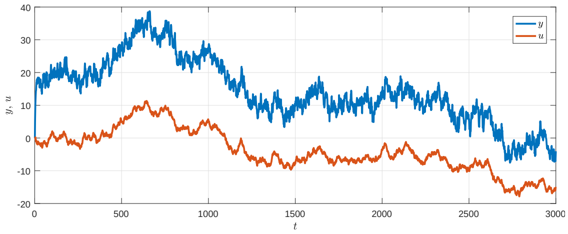

Figure 5: Simulated example – input and output.

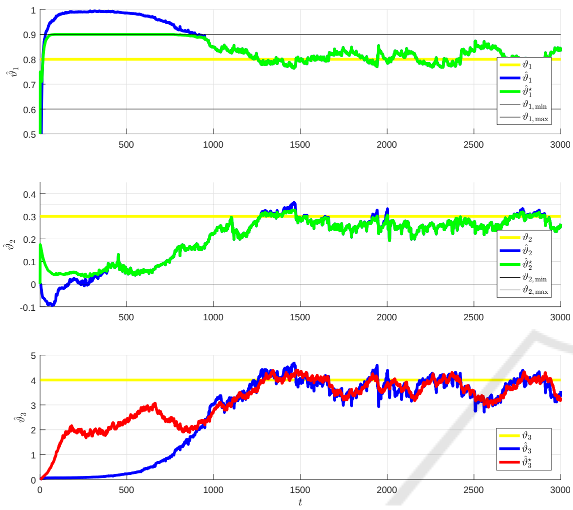

Figure 6: Simulated example – behaviour of unbounded (blue) and bounded (green) parameter estimates. Estimated parameter of the reduced model is plotted in red.

4.2.4 Real Data Example

The example is based on the real data set which was used in (Ettler and Kárný, 2010). System output y is represented by deviation of the output strip thickness during the process of cold rolling. Its behaviour is being approximated by the model

$$y(k) = \vartheta_1 u(k-1) + \vartheta_2 v(k-1) + \vartheta_3 + e_y(k) \; , \quad (60)$$

where index k means sample number while the sampling is triggered by the movement of the rolled strip and $\Delta k \sim 0.08$ m of the strip length. Control signal u corresponds to so-called uncompensated rolling gap of the rolling mill and the measured disturbance v is represented by the nonlinear function of the rolling force. The model is based on the *gaugemeter principle*, see, e.g. (Ettler and Andrýsek, 2007) for details.

Fig. 7 shows undesirable variations of measured output thickness in the interval $k \in (3300, 4200)$ which was caused by dirt influencing the contact thickness measurement.

Situation depicted in Fig. 7 caused a temporary discrepancy between both sides of the model (60) which may drive unbounded parameter estimates out of their reasonable ranges.

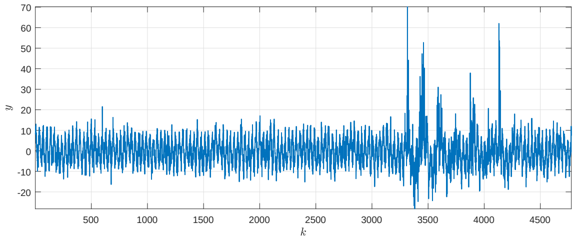

Figure 7: Real data example – the measured output.

Based on the knowledge of the process, it was possible to determine boundaries for the first two model parameters:

$$\begin{array}{rcrcrcr} \underline{\vartheta}_1 & = & -1.0 & \quad & \overline{\vartheta}_1 & = & -0.02 \\ \underline{\vartheta}_2 & = & -100.0 & & \overline{\vartheta}_2 & = & 0.00 \\ n & = & 4700 & & & & \end{array} \quad (61)$$

Behaviour of parameter estimates is shown in Fig. 8. The second unbounded parameter (in blue) needs relatively long time $k \in \langle 1, 700 \rangle$ to reach the allowed range, which can be explained by a moderate excitation of the model because selected data come from the middle of the rolled strip. Concerning the bounded estimation, the situation is balanced by the parameter of the reduced model (in red) in the third graph.

Note: Real on-line identification starts with the beginning of rolling when the system is excited enough. Here the off-line identification starts in the steady-state to illustrate behaviour of the estimator under unfavorable conditions.

As a consequence of the measurement error during $k \in (3300, 4200)$, the blue unbounded estimates $\hat{\vartheta}_1$, $\hat{\vartheta}_2$ exceeded their limits. The bounded estimator coped with the problem reasonably which is illustrated by the behaviour of the green bounded estimates. Again, the red parameter of the reduced model deviated from its unbounded version to minimize the prediction error.

5 CONCLUSIONS

The paper deals with the *realistic* estimation of model parameters which takes into account limitation on model parameters existing in real applications.

In general, limitation of a random variable with normal distribution $\mathcal{N}$ is described by the introduced heterogenous probability distribution $\mathcal{D}$. Relations between mean and variance of both distributions can be adequately approximated by cdf and pdf of the generalized normal distribution $\mathcal{G}$, respectively. After

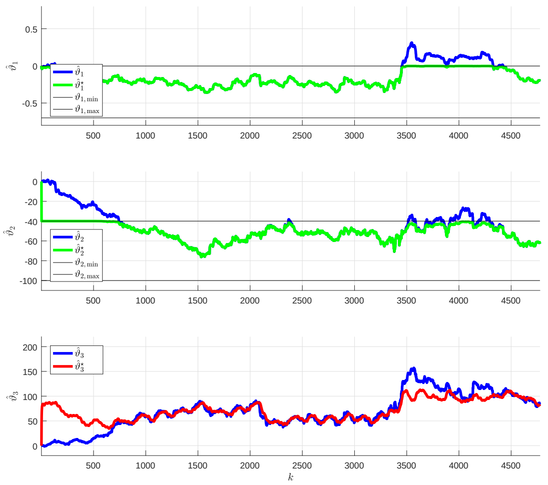

Figure 8: Real data example – behaviour of unbounded (blue) and bounded (green) parameter estimates. The estimated parameter of the reduced model is plotted in red.

determination of rules for construction of the fictive $\mathcal{G}$ it was possible to introduce the limiting functions for the mentioned statistical moments and integrate them into recursive algorithm of bounded parameter estimation. In addition, parallel estimation of the *full* and *reduced* models enable to minimize the prediction error in each estimation step.

Behaviour of the estimator was illustrated on simulated data and then on real data taken from a cold rolling mill.

REFERENCES

Åström, K. J. and Kumar, P. (2014). Control: A perspective. *Automatica*, (50):3–43.

Benavoli, A., Chisci, L., Farina, A., Ortenzi, L., and Zappa, G. (2006). Hard-constrained vs. soft-constrained parameter estimation. *IEEE Transactions on aerospace and electronic systems*, 42(4):1224 – 1239.

Ettler, P. and Andrýsek, J. (2007). Mixing models to improve gauge prediction for cold rolling mills. In *Proceedings of the 12th IFAC Symposium on Automation in Mining, Mineral and Metal Processing*, Québec, Canada.

Ettler, P. and Kárný, M. (2010). Parallel estimation respecting constraints of parametric models of cold rolling. In *Proceedings of the 13th IFAC Symposium on Automation in Mineral, Mining and Metal Processing (IFAC MMM 2010)*, pages 63–68, Cape Town, South Africa.

Kárný, M. (1982). Recursive parameter estimation of regression model when the interval of possible values is given. *Kybernetika*, 18(2):164–178.

Kárný, M., Böhm, J., Guy, T., Jirsa, L., Nagy, I., Nedoma, P., and Tesař, L. (2005). *Optimized Bayesian Dynamic Advising: Theory and Algorithms*. Springer, London.

Kopylev, L. (2012). Constrained parameters in applications: Review of issues and approaches. *ISRN Biomathematics*, 2012:Article ID 872956.

Kulhavý, R. and Zarrop, M. B. (1993). On a general concept of forgetting. *International Journal of Control*, 58(4):905–924.

Mandelkern, M. (2002). Setting confidence intervals for bounded parameters. *Statistical Science*, 17(2):194–172.

Milanese, M., Norton, J., Piet-Lahanier, H., and (Eds.), E. W. (1996). *Bounding Approaches to System Identification*. Springer.

Murakami, K. and Seborg, D. E. (2000). Constrained parameter estimation with applications to blending operations. *Journal of Process Control*, 10:195–202.

Norton, J. P. (1987). Identification and application of bounded-parameter models. *Automatica*, 23(4):497–507.

Peterka, V. (1981). *Bayesian Approach to System Identification In P. Eykhoff (Ed.) Trends and Progress in System Identification*. Pergamon Press, Eindhoven, Netherlands.

Toulias, T. L. and Kitsos, C. P. (2014). On the properties of the generalized normal distribution. *Discussiones Mathematicae Probability and Statistics*, 34(1-2):3549.

APPENDIX

Generalized Normal Distribution

Symmetric version of the generalized normal distribution $\mathcal{G}$ is defined by 3 parameters: $\mu_{\mathcal{G}}$ (location), α (scale) and β (shape).

Pdf of $\mathcal{G}$ is given by

$$f_{\mathcal{G}}(x|\mu_{\mathcal{G}},\alpha,\beta) = \frac{\beta}{2\alpha\Gamma(1/\beta)} \exp\left\{-\left|\frac{x-\mu_{\mathcal{G}}}{\alpha}\right|^{\beta}\right\}, \quad (62)$$

where $\alpha > 0$, $\beta > 0$ and Γ denotes the gamma function

$$\Gamma(x) = \int_0^{\infty} t^{x-1}\exp(-t)dt\ . \quad (63)$$

For $\beta = 2$, $\mathcal{G}$ coincides with the normal distribution $\mathcal{N}(\mu_{\mathcal{G}}, \frac{\alpha^2}{2})$. For $\beta \to \infty$, $\mathcal{G}$ converges pointwise to uniform density on $(\mu_{\mathcal{G}}-\alpha, \mu_{\mathcal{G}}+\alpha)$.

Cdf of $\mathcal{G}$ is given by

$$F_{\mathcal{G}}(x|\mu_{\mathcal{G}},\alpha,\beta) = \quad (64)$$
$$\frac{1}{2}\left[1+\frac{\mathrm{sgn}(x-\mu_{\mathcal{G}})}{\Gamma(1/\beta)}\,\gamma\left(1/\beta, \left|\frac{x-\mu_{\mathcal{G}}}{\alpha}\right|^{\beta}\right)\right],$$

where γ means the lower incomplete gamma function

$$\gamma(x,x_0) = \int_0^{x_0} t^{x-1}\exp(-t)dt\ . \quad (65)$$

Remaining properties of $\mathcal{G}$ can be found for example in (Toulias and Kitsos, 2014).